AFRL-RH-WP-TR-2010-0110

VALIDATION OF THE AGGREGATE WAKEFULNESS AND READINESS ESTIMATOR (AWARE) USING ON-THE-JOB SECURITY FORCE PERSONNEL

J. Lynn Caldwell   Christienne Ruth

Biosciences and Performance Division
Vulnerability Analysis Branch

Margaret Funke   Lindsey McIntire

InfoSciTex
Dayton OH

William F. Storm

Wylie Laboratories
Brooks City Base, San Antonio TX

Julia N. Sundstrom

Lackland Air Force Base
San Antonio TX

August 2010

Approved for public release, distribution is unlimited.

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING,
HUMAN EFFECTIVENESS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2010-0110 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//	//SIGNED//
Suzanne Smith, Work Unit Manager Vulnerability Analysis Branch	Mark M. Hoffman, Deputy Chief Biosciences and Performance Division Human Effectiveness Directorate 711th Human Performance Wing Air Force Research Laboratory

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS**.			
1. REPORT DATE ***(DD-MM-YY)*** 19-08-2010	**2. REPORT TYPE** Interim		**3. DATES COVERED** ***(From - To)*** July 7 2009- July 6 2010
4. TITLE AND SUBTITLE Validation of the Aggregate Wakefulness And Readiness Estimator(AWARE) using On-the-Job Security Forces Personnel			**5a. CONTRACT NUMBER** IN-HOUSE
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER 62202F
6. AUTHOR(S) *J. Lynn Caldwell, Christienne Ruth **Margaret Funke, Lindsey McIntire ***William F. Storm ****Julia Sundstrom Tom Beltran			**5d. PROJECT NUMBER** 7184
			5e. TASK NUMBER 02
			5f. WORK UNIT NUMBER 71840223
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **InfoSciTex, 4027 Colonel Glenn Highway, Suite 210, Dayton OH ***Wiley Laboratories, San Antonio, TX ****Air Force Security Forces Center, Lackland AFB, TX			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) *Air Force Materiel Command Air Force Research Laboratory 711th Human Performance Wing Human Effectiveness Directorate Biosciences and Performance Division Vulnerability Analysis Branch Wright-Patterson AFB, OH 45433-7947			**10. SPONSORING/MONITORING AGENCY ACRONYM(S)** 711 HPW/RHPA
			11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S) AFRL-RH-WP-TR-2010-0110
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			
13. SUPPLEMENTARY NOTES 88ABW cleared on 1 Sep 2010, 88ABW-2010-4772			
14. ABSTRACT The present study sought to document the work/rest schedule as well as subjective and objective fatigue levels of Air Force security forces personnel and apply these data to a performance prediction model. Participants were asked to complete a sleep diary each morning and mood and performance tests every 2 hours during their work days, and upon rising, 10 hours into the day, and then prior to bedtime on days off. Due to the participation variability among shifts and bases, the results from the study do not allow direct comparison of shift schedules nor time on shift as hoped. However, the data were able to show that as time on shift increased, both subjective and objective fatigue generally increased, particularly on the 12-hr shifts.			
15. SUBJECT TERMS Security Forces; Shift work; Fatigue			

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT:**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** (Monitor)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			J. Lynn Caldwell
Unclassified	Unclassified	Unclassified	SAR	33	**19b. TELEPHONE NUMBER** *(Include Area Code)* N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

THIS PAGE IS INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENT

We would like thank the commanders involved in the study: Lieutenant Colonel Edward Schneider at Lackland AFB, and Major David Harris and Senior Master Sergeant Stephanie Liles at Wright Patterson AFB. Without their support, this study would not have been possible. We would also like to thank Senior Technical Sergeant Chris Meyers for contributing to preparations for the study. And finally, we would like to thank all those airmen who volunteered to participate in the study and gave many hours of their time to contribute to the data base; without their participation, this study would not have succeeded.

In addition to the study participants, we would like to express our sincere gratitude to the people at MTS Technologies, Inc., and the Institute of Behavioral Resources, Inc. (IBR) for their funding and guidance with this study. Without their efforts, this study would not have come to fruition.

THIS PAGE IS INTENTIONALLY LEFT BLANK

SUMMARY

The present study sought to document the work/rest schedules as well as subjective and objective fatigue levels of Air Force security forces personnel and apply these data to a performance prediction model. Participants in the study included personnel from Wright-Patterson AFB (WPAFB) and Lackland AFB (LAFB), representing two different shift schedules. WPAFB uses 3, 8-hr shifts while LAFB AFB uses 2, 12-hr shifts. Participants were asked to complete a sleep diary each morning and mood and performance tests every 2 hours during their work days, and upon rising, 10 hours into the day, and then prior to bedtime on days off.

Due to the participation variability among shifts and bases, the results from the study do not allow direct comparison of shift schedules nor time on shift as hoped. However, the data were able to show that as time on shift increased, both subjective and objective fatigue generally increased. Caution is used in interpreting the performance data from the Psychomotor Vigilance Test (PVT). The PVT is generally performed in a quiet environment while the participant is seated. The present study presented the PVT on a Personal Digital Assistant (PDA) with participants taking it in wherever they were at the time the test was scheduled, which may have been a noisy environment and/or while seated or standing.

A subset of the data from LAFB was submitted to MTS Technologies, the contractor for this effort. The data will be used to determine the validation of the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) model as used in the Aggregate Wakefulness And Readiness Estimator (AWARE) software.

In addition to validation of the AWARE software, summarized data were presented to the two commanders of the squadrons who participated in the study. While the information will not be used to alter shift schedules, it provided each commander with an idea of how alertness and performance declines across the day, particularly during the night shifts and on the longer, 12-hr shifts.

1.0 INTRODUCTION

In the U.S. military, personnel are often required to engage in 24/7 operations. Security Forces personnel are a prime example of military units which must be staffed around the clock, both in peacetime and during war. Common schedules for security forces include 3, 8-hr shifts, or 2, 12-hr shifts, which are either rotating or fixed. Shifts usually involve at least 2 hours of additional time over the scheduled time to allow personnel to check in/out weapons and complete paperwork. As a result, prolonged work bouts are common and shorter-than-normal sleep periods are unavoidable, resulting in mental and/or physical fatigue which can impede operational readiness. It is well established that prolonged wakefulness and the resulting cumulative sleep debt increase the likelihood that personnel will briefly (and uncontrollably) nod off on the job, even during demanding tasks (Caldwell, Caldwell, and Schmidt, 2008). The longer amount of time personnel remain awake, the more likely “sleep attacks” become. Sleepiness takes a heavy toll on reaction time, motivation, attention, memory, endurance, and judgment as well (Balkin, Rupp, Picchioni, and Wesensten, 2008).

In addition to long hours, working during the night hours creates performance and alertness problems beyond those associated with daytime work hours. Night work is associated with impaired performance, disturbed sleep, and an increased rate of accidents caused by circadian desynchrony and cumulative sleep deprivation (Akerstedt, 1995; Torsvall et al., 1989). Coleman and Dement (1986) found that 53 percent of shift workers report falling asleep during the night shift at least once each week, while only 8 percent of permanent day workers report falling asleep. Another study reported 20 percent of shift workers fell asleep during the night shift, while none fell asleep during the evening or afternoon shifts (Torsvall et al., 1987). Additionally, degraded performance related to night work has been reported to contribute to transportation accidents (Dinges, 1994; Philip and Akerstedt, 2006) and major industrial disasters such as the Exxon-Valdez oil spill (National Transportation Safety Bureau, 1990) and the Chernobyl meltdown (U.S. Nuclear Regulatory Committee, 1987). Even moderate reduction in sleep can affect human perception, cognitive speed, and decision making (Belenky et al., 2003; Van Dongen et al., 2003).

One strategy for countering the effects of reduced sleep, due to long work hours or working during the nighttime hours, on human performance is more optimal scheduling of work and sleep. The Department of Defense (DoD) sponsored development of a scientific model of sleep, wakefulness and performance known as the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) model (Hursh, 2004). The SAFTE model has been validated extensively with laboratory data and has been used in the construction of a software tool called the Aggregate Wakefulness And Readiness Estimator (AWARE). Incorporating the SAFTE model into AWARE allows users to input up to a month of specific work/sleep schedules to determine the degree of general performance impairments that can be expected with a projected schedule. This tool provides an efficient way to model the impact of proposed schedule revisions. However, revised versions of SAFTE need validation, and user assessment of the AWARE software interface is needed to ensure its accuracy, define its limitations in operational settings, and determine its ease of use and clarity of output. Therefore, a study was designed to collect sleep/wake and performance data from working participants after which, these data would be input into the AWARE software to validate it against real-world data.

2.0 METHOD, ASSUMPTION AND PROCEDURES

2.1 Equipment

Cognitive tests and questionnaires were presented on a personal digital assistant (PDA), also known as a palmtop computer. The questionnaires included on the PDA were the Visual Analogue Scale, the Profile of Mood States Brief, and a sleep diary (collects accounts of bedtime, wake time, and subjective sleep quality). The Psychomotor Vigilance Task, a performance measure, was included as well. Objective rest/activity data were collected and stored on a wrist activity monitor (Actiwatch® from Mini-Mitter/Respironics).

2.1.1 Visual Analogue Scale (VAS). The Visual Analogue Scale (VAS) is a questionnaire that measures subjective attitudes, including sleepiness and alertness. The nine adjectives used to help the participant define how they were feeling "right now," were: alert/able to concentrate, energetic, confident, irritable, jittery/nervous, sleepy, and talkative. Each adjective was presented under a line that was defined at one end by "not at all" and by "extremely" on the other. To complete the questionnaire, the participant put a mark on the line that corresponded to how they felt along the continuum for each adjective.

2.1.2 Profile of Mood States Brief (POMS-B). The Profile of Mood States Brief (POMS-B) is a 30-item questionnaire that measures mood using 6 categories: tension-anxiety, depression-dejection, anger-hostility, vigor-activity, fatigue-inertia, and confusion-bewilderment. Participants rated their feelings about each item (example items: Tense, Lively, Uneasy) on a scale of 1 to 5, with 1 being "not at all" and 5 being "extremely." The overall scores for each of the five categories were totaled, resulting in five factor scores. Each of the factor scores, except for the vigor-activity score, was added together, then the vigor-activity factor score was subtracted from this total producing the general mood disturbance score.

2.1.3 Sleep diary. The sleep diary questionnaire focused on questions such as the number of times the participant woke, how rested they felt, what may have contributed to his/her restlessness (if any was experienced), and any sleep practices or sleep aids which might have been used. Of the items in the test battery, the sleep diary was always the first thing that participants completed upon waking.

2.1.4 Psychomotor Vigilance Task (PVT). The Psychomotor Vigilance Task (PVT) is a simple reaction time test. In this study, the test length was 5 minutes during which the participant waited for the stimulus to appear on the touch screen of the PDA and then pressed the screen with his/her thumb (using the dominant hand) as fast as they could to indicate a response, thus recording his/her reaction time. The PDA visually displayed numbers counted up by milliseconds, presented up to 1 minute (60,000 msec), allowing the participant to respond. The number continued to count up until the participant responded or the trial timed out. The inter-stimulus interval of the stimuli varied randomly from 2 to 12 seconds. Metrics obtained from this test were reaction time (RT) and lapses, defined as a response which occurred later than 500 msec (or not at all).

2.1.5 Wrist activity monitors (WAM).

The wrist monitors (Actiwatch® by MiniMitter/Respironics) were battery-powered devices about the size of a wrist watch that were used to track sleep/activity rhythms. The participants each wore one for the length of the study.

2.2 Participants

Participants were security forces personnel recruited from two U.S. Air Force bases (Wright-Patterson Air Force Base (WPAFB), OH, and Lackland Air Force Base (LAFB), TX). There were no gender restrictions; the qualifying criteria were active duty military age 18 or older working as an Air Force Security Forces member. There were 26 participants recruited from the WPAFB, of whom 23 were men and 3 were women. At LAFB, 36 individuals volunteered of whom 29 were men and 7 were women. Compensation was not provided to any participant.

2.2.1 Duration. At LAFB, participants were asked to participate in the study for 11 days and to return their materials on day 12. At WPAFB, participants were asked to participate for 14 days, and their materials were collected on day 15. In all cases, the first day of data collected was not included in the data analysis because the researchers considered the time to be practice.

2.3 Description of study

2.3.1 Data Collection. At WPAFB, the Security Forces schedules included 3, 8-hr shifts. The first shift began at 0700 and ended at 1500, the second shift began at 1500 and ended at 2300, and the third shift began at 2300 and ended at 0700. In contrast, the schedules at LAFB are divided into 2, 12-hr shifts. Those on the Alpha and Bravo teams report to work at 0515 and finish at 1715. The Charlie and Delta teams work from 1715 until 0515. The Alpha team is always paired with the Charlie team, and the Bravo team always works with the Delta team. On days that the Alpha and Charlie teams had off, the Bravo and Delta teams were on duty and vice versa.

Each participant was asked to participate for a minimum of one work week (work days plus rest days). They were asked to wear the Actiwatch® the entire length of the study in order to obtain estimates of activity and sleep during the data collection period. In addition, participants also carried their PDAs on which they took the tests and questionnaires. At a minimum, participants performed the test battery upon arising from sleep, upon reporting to duty, at the end of the work shift, and just before bedtime. On rest days, participants performed the test battery upon arising from sleep, 10 hours post-awakening, and just before bedtime. Participants at WPAFB were asked to take additional tests at 2-hr intervals during the work period when possible, and participants at LAFB were asked to take additional tests at 4-hr intervals during the work period when possible. The participants at LAFB were asked to take fewer tests because their shifts were longer. The PDAs were programmed to vibrate/alarm at the desired time of testing. The testing schedule is shown in Table 1.

Table 1. Schedule of test time

Time	Duty Day 1	Duty Day 2	Duty Day 3	Duty Day 4	Duty Day 5	Rest Day 1	Rest Day 2
Wake-up	*	*	*	*	*	*	*
Report for duty	*	*	*	*	*		
Duty + 2 hours							
Duty + 4 hours							
Duty + 6 hours							
Duty + 8 hours							
Duty + 10 hours							
Duty + 12 hours							
End of Shift	*	*	*	*	*		
Wake + 10 hours						*	*
Bedtime	*	*	*	*	*	*	*

* indicates required testing

Each time the participants took a test, the tasks in the battery and the order of those tasks were always the same. The participant completed the POMS first, then the VAS, and finished with the 5-min PVT. The only time that the battery varied was to include the Sleep Diary which was presented first when participants initiated the post-sleep test just after the PVT was completed.

The data from the performance measures will also be input into the SAFTE model to determine how well the model predicts the performance curves from all individuals. A modification to the SAFTE model was created, and these data were used to determine how well the model predicts the performance curves from all individuals. Though the model has been tested using laboratory data as the input, this study provided real-world data with which to verify the model.

3.0 RESULTS AND DISCUSSIONS

Since this study had little day-to-day involvement of the experimenters, its success relied heavily on the participants' willingness to update their diaries (in the PDAs) to reflect their activities and to continue to take the battery of tests at the designated times. For example, if each of the 36 participants at LAFB had taken every test when it became available over the course of the 10-day study, they would have generated data for 1,800 tests. It is unfortunate, therefore, that although many of the participants had strong beginnings, consistent participation began to decline after a few days in many cases. Additionally, participants sometimes skipped tests and/or took them at inappropriate times, which created a situation in which the data from many tests had to be discarded. In order to separate the data, the researchers compared the times that the tests were taken with the information in the participants' diaries and with their work schedules. The sleep/wake times recorded by the participants' Actiwatch® were also referenced in deciding which data to use in the analysis. The researchers concluded that any test that was taken within +/-1 hour of the scheduled test time was considered viable. Therefore, if a participant took a

Post-sleep test 45 minutes after waking (as determined by the Actiwatch® data), the data from the test were kept.

After separating out the usable data, the researchers determined they had data from only a fraction of the number of tests expected, and much of that data came from tests during the workday. As a result, they narrowed their focus to tests taken during the workday (excluding post-sleep and before sleep tests). It was also necessary that these workday tests were from days that began with a pre-shift test, so that there was a baseline with which to make comparisons. Lastly, a "day" of data was further defined as any workday where the participant took a pre-shift test and then at least one more test during his/her shift (post-shift tests included). Ultimately, the data included participation from 43 days (146 usable tests) that were generated from 25 participants (18 men, 7 women) at LAFB. Similarly, there were 24 days (and 98 usable tests) of data collected from 12 (11 men, 1 woman) participants at WPAFB.

Additionally, because there were several days of usable data from some of these participants and only one day of usable data from others, the data from each day were treated as independent data points, regardless of which participant contributed it. Therefore, the *n*'s reported in the data analysis reflect the number of tests that contributed to those means and standard deviations rather than the number of participants who contributed data. Only descriptive statistics are given; no statistical analysis of time on shift, shift differences, or base comparisons is calculated, and therefore, no error bars are included in the graphs.

3.1 Wright Patterson Air Force Base (WPAFB)

The number of participants for this data set was limited. As mentioned above, the means for the reported data are based on the number of tests completed, not the number of participants. Some of the data points are contributed by the same individual over several days of the study. The number of data points analyzed in each data set is shown in Table 2.

Table 2. Number of responses per work shift and time of testing for each variable

Dataset	Shift	Preshift	+2	+4	+6	+8	Postshift
VAS POMS-B PVT	1st	8	8	7	5	3	5
	2nd	8	8	7	8	5	6
	3rd	3	3	2	3	2	2

Dataset	Shift	n
Sleep Diary	1st	7
	2nd	11
	3rd	13
Wrist monitors (WAM)	1st	61
	2nd	56
	3rd	33

3.1.1 Visual Analogue Scale (VAS). Overall, participants from the three shifts at WPAFB had high to moderately-high levels of self-reported alertness, energy, and confidence. Their anxiety and nervousness scores were low, and their levels of self-reported irritability, talkativeness, and sleepiness were low to moderately-low.

Due to the low *n*'s in the data set, direct comparisons among shifts cannot be made statistically; however, visual comparisons reveal interesting trends. The levels of alertness, energy, confidence, and sleepiness were all very similar among respondents in the three shifts. However, the first shift recorded having higher levels of talkativeness than the other two shifts. The levels of anxiety reported in the first and third shifts were also comparable, but the reported levels were slightly higher during the second shift. A similar pattern is also observable in the jittery/nervousness scores.

Of the three shifts, the participants in the first shift reported the highest mean levels of irritability at baseline (pre-shift test). By 0900 (2 hours into shift), there was a noticeable decrease in irritability. Irritability decreased even more at 1100 (4 hours into shift), and it stayed low for the rest of the day. At 1100 (4 hours into shift), along with decreased irritability, the participants in the first shift also reported higher mean energy and anxiety levels. Although, those means returned to baseline levels by 1300 (6 hours into shift), their means scores for alertness and talkativeness climbed. At the end of the day, at 1500 (8 hours into shift), the participants were more alert, more energetic, and more confident than they were when the day began. They were still as talkative as they were at 1300 as well.

The second shift reported feeling more energetic at 1700 (2 hours into shift) than they did at baseline (pre-shift test). However, by 1900 (4 hours into shift) the mean energy level was back to baseline. These participants indicated a loss of confidence and greater sleepiness 6 hours into the shift (2100). At the end of their shift (2300), the group's mean for sleepiness was still as low as at 2100, and the participants reported having less energy than at baseline too. Post-shift tests revealed that this group felt more anxious and jittery/nervous, and less energetic than they did at baseline. Their levels of sleepiness, at the time of the post-shift test, were still higher than baseline.

Unlike the second shift, the participants in the third shift reported a higher group mean for irritableness and sleepiness at 0100 (2 hours into shift) than they did at baseline. At 0300 (4 hours into shift), their mean score for irritableness was still high, but these participants reported feeling more nervous and more talkative as well. The mean scores for each of those levels returned to baseline at 0500 (6 hours into shift). By 0700, the end of their shift, these individual reported being less alert and sleepier than they were at baseline. Their mean alertness scores returned to baseline levels during the post-shift test, but their reported sleepiness remained as high as it had been at 0700.

The time course of the various scales from the VAS is shown in Figure 1 below.

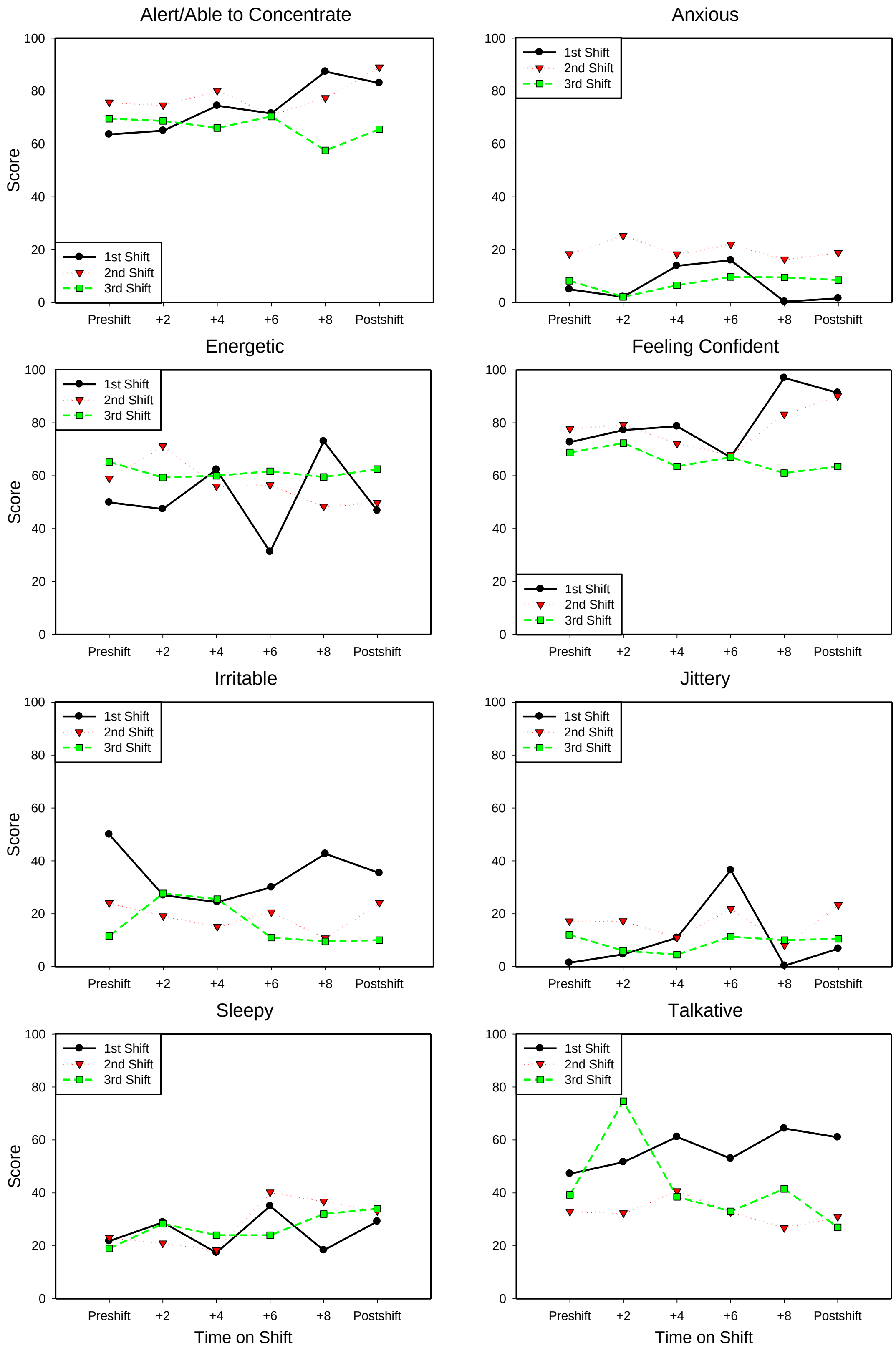

Figure 1. VAS scores by shift and time of day

3.1.2 Profile of Mood State-Brief (POMS-B)

*3.1.2.1 Mood Disturbance Scores***.** At WPAFB, the mean Mood Disturbance Scores for the participants in all three shifts stayed close to their baseline scores throughout each of their workdays, with the third shift showing a slight decline towards the last 2 hours of their work period. The data are represented in Figure 2 below.

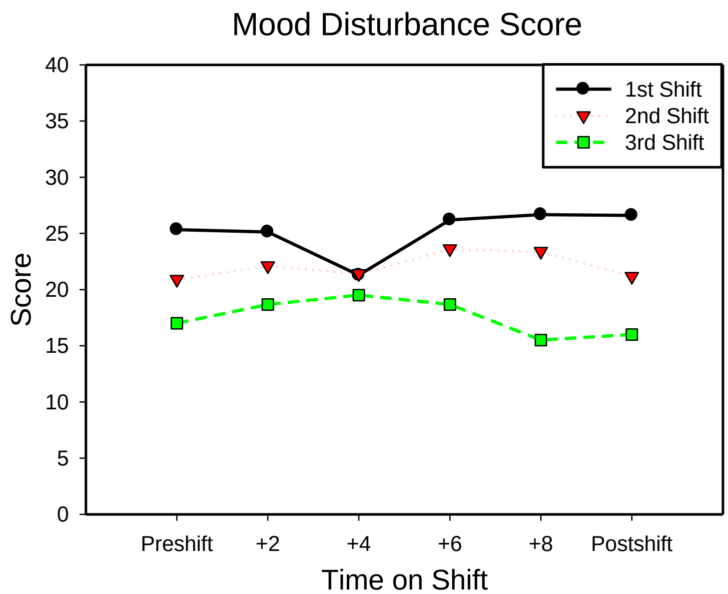

Figure 2. POMS Mood Disturbance Score by shift and time of day

*3.1.2.2 Factor scores***.** At WPAFB, there were very few large changes in the mean factor scores reported throughout the day. The only notable change that occurred during the first shift was in the reported anger/hostility scores. When the participants had been at work for 4 hours (1100), their mean anger/hostility score dropped below baseline and stayed below the baseline level throughout the rest of the shift. During the second shift, the mean score for vigor/activity dropped below baseline at 2300 (8 hours into the shift) and fatigue/inertia rose above baseline. With the post-shift tests, the vigor/activity score fell even further below baseline, but the mean score for the fatigue/inertia returned to near-baseline levels. The only remarkable change that occurred during the third shift was the mean score for anger/hostility dropped below the baseline level during the post-shift tests (after 0700). The data are shown in Figure 3 below.

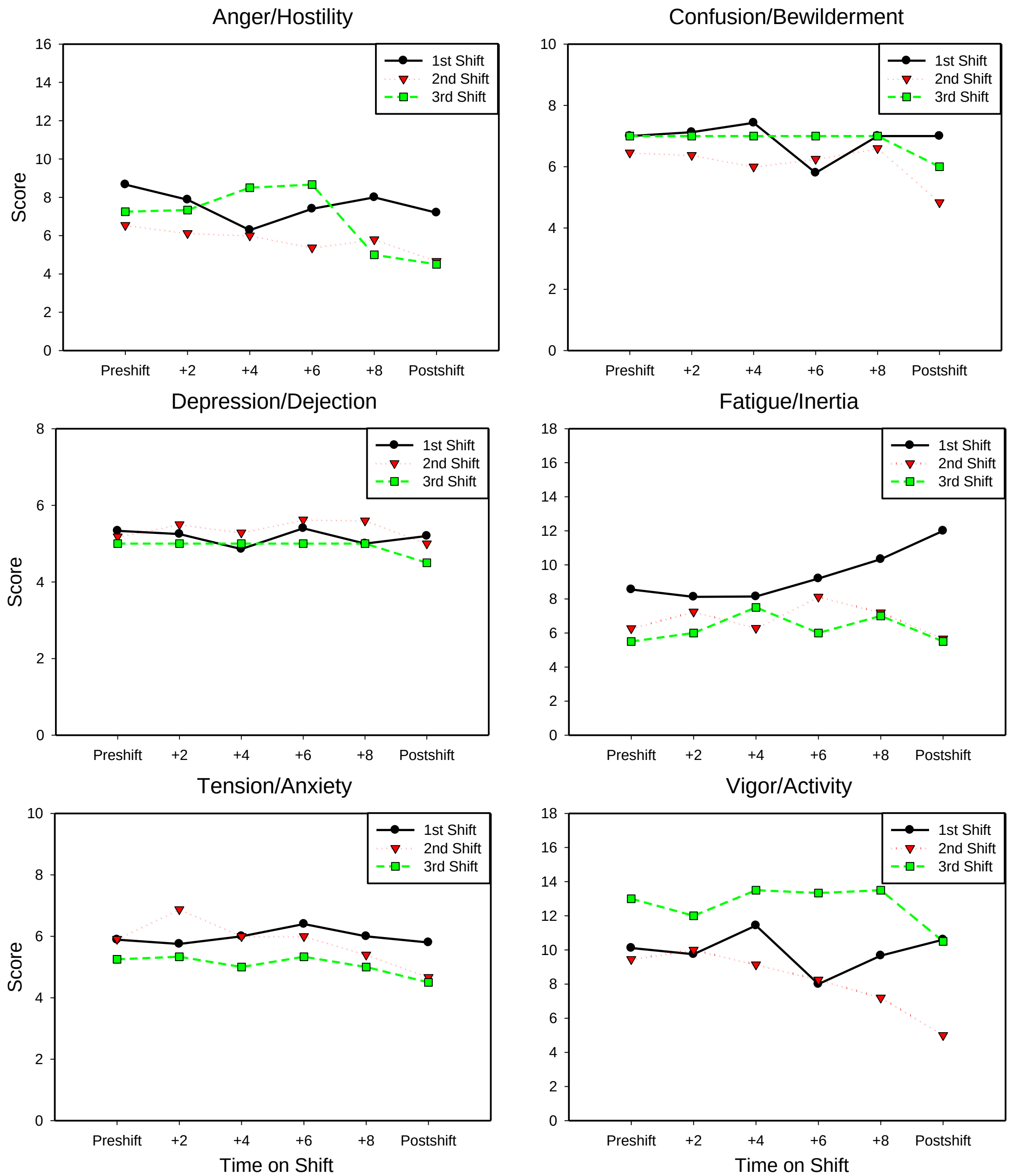

Figure 3. POMS factor scores by shift and time of day

3.1.3 Sleep diary. For the sleep diaries at both WPAFB and LAFB, the *n* used for each shift actually reflects the total number of sleep-periods that occurred on workdays from which usable data were pulled, instead of the number of individuals participating.

The participants from the second shift indicated that they slept quite a bit better than the other two groups. However, the second and third shifts responded they had less trouble falling asleep and slept deeper than the participants in the first shift. On average, it also took these two groups less time to fall asleep than the participants from the first shift. When asked how rested they felt, the first shift's mean score was quite a bit lower than the other two shifts, indicating that they felt the least rested. The participants from the first shift also indicated they awakened more during their sleep period than the participants in the other two shifts, a possible reason why they indicated feeling less rested than the other participants. The data are shown in Figure 4.

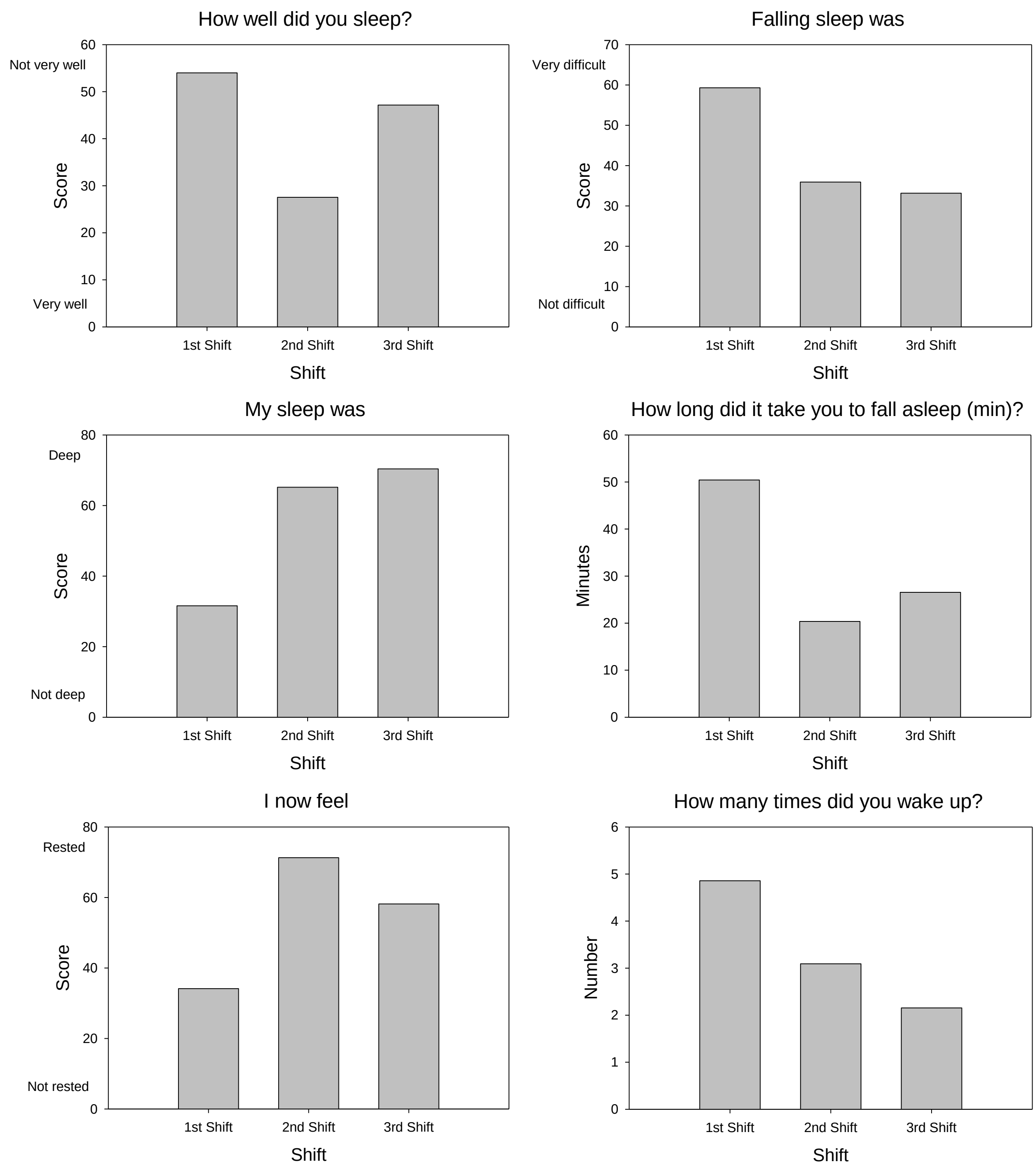

Figure 4. Sleep questionnaire responses concerning sleep quality by shift

None of participants reported using relaxation techniques or alcohol to help them sleep, but a small percentage in each shift reported using good sleep habits to help them sleep. Of the three groups, only a few participants in second shift reported using herbal remedies, over-the-counter medicines, or other techniques to help them sleep. Figure 5 shows the data from these questions.

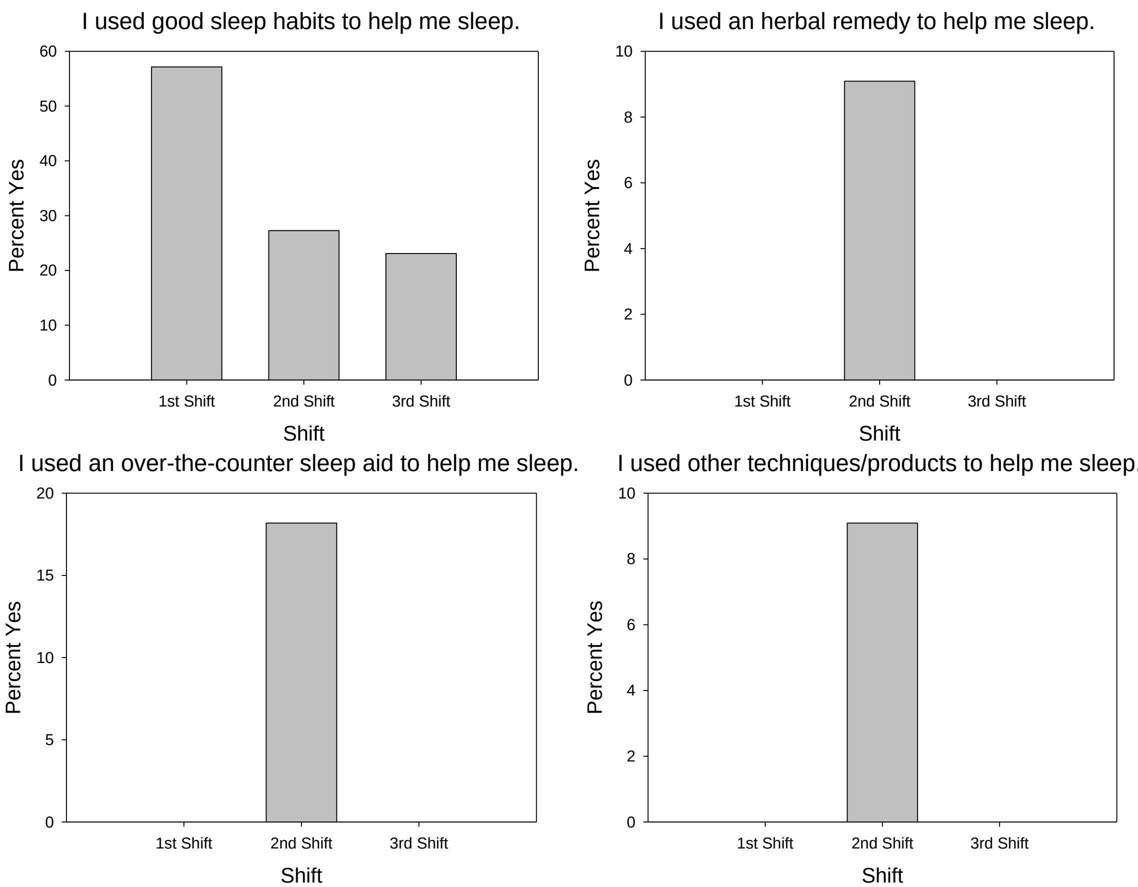

Figure 5. Sleep questionnaire responses by shift concerning aids to obtain sleep

A few people from each shift reported that external physical and internal physical discomforts affected the quality of their sleep; there were also a few reports from each shift that not being sleepy affected the quality of their sleep. In each case, there were more people indicating that that was true in the first shift than in the other two shifts. Additionally, one person from the first group and one from the third reported that a disruptive environment affected the quality of their sleep. A small percentage of people in each shift also reported that personal stress/anxiety/worry affected the quality of their sleep, with a higher percentage of people in the first and second groups reporting this to be true than in the third group. Finally, a few participants indicated other factors affected the quality of sleep, more on the third shift than the other two shifts. Figure 6 shows the data for each shift for these questions.

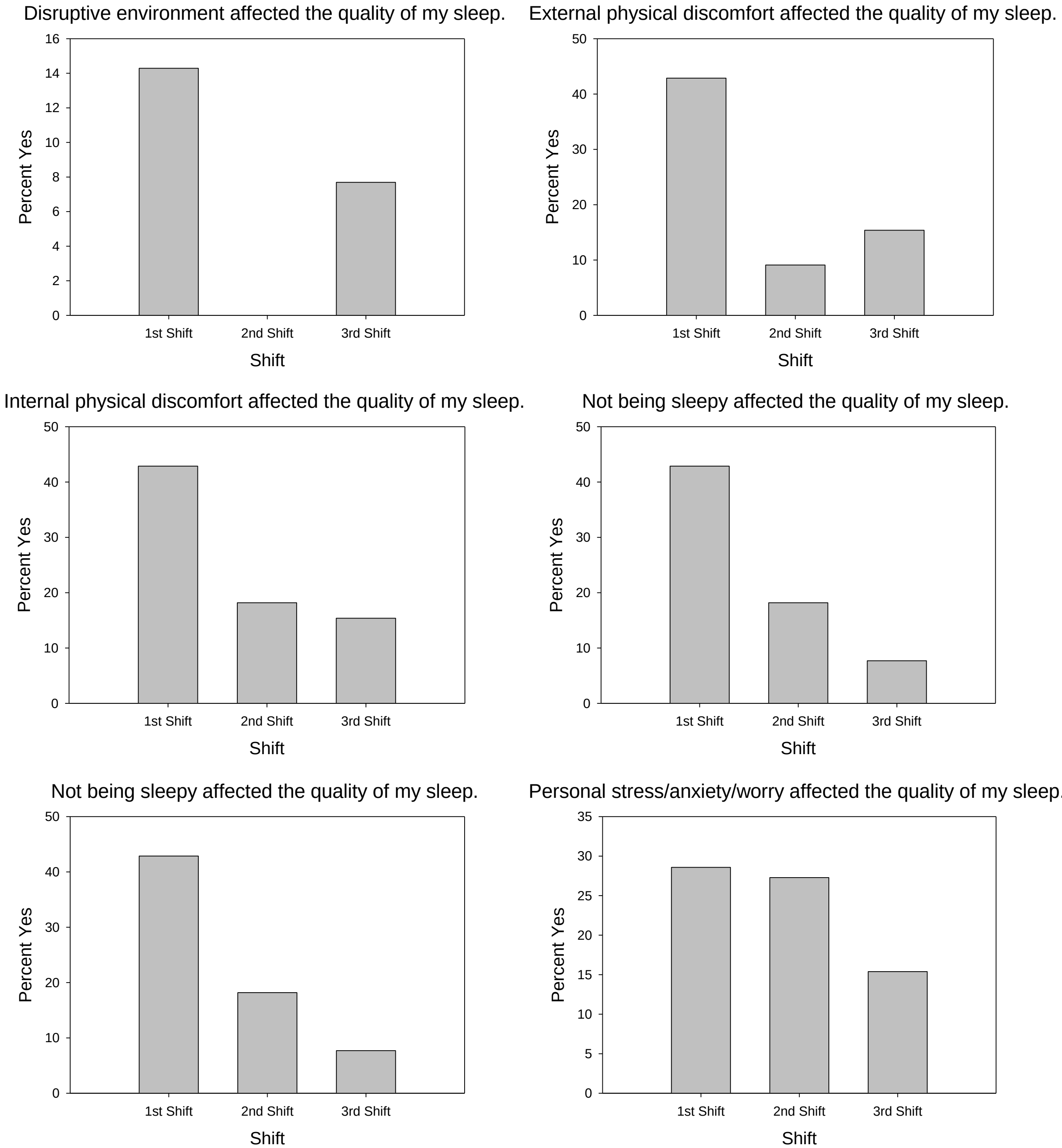

Figure 6. Sleep questionnaire responses by shift concerning factors affecting sleep quality

3.1.4 Psychomotor Vigilance Task (PVT). The data collected with the Psychomotor Vigilance Task (PVT) are reaction times and number of lapses. Generally, reaction times greater than 500 msec (lapses) are often associated with fatigue or sleepiness. They can also occur when a participant is giving less than his/her full attention to the task, which was a possibility in this study since the participants were taking the tests while on duty.

The average reaction times each of the shifts showed varying times with no real consistent pattern as did the number of lapses per shift. Perhaps the confound of performing the PVT while also needing to perform one's job does not allow the attention needed to correctly perform this task. The data for each shift are presented in Figure 7 below.

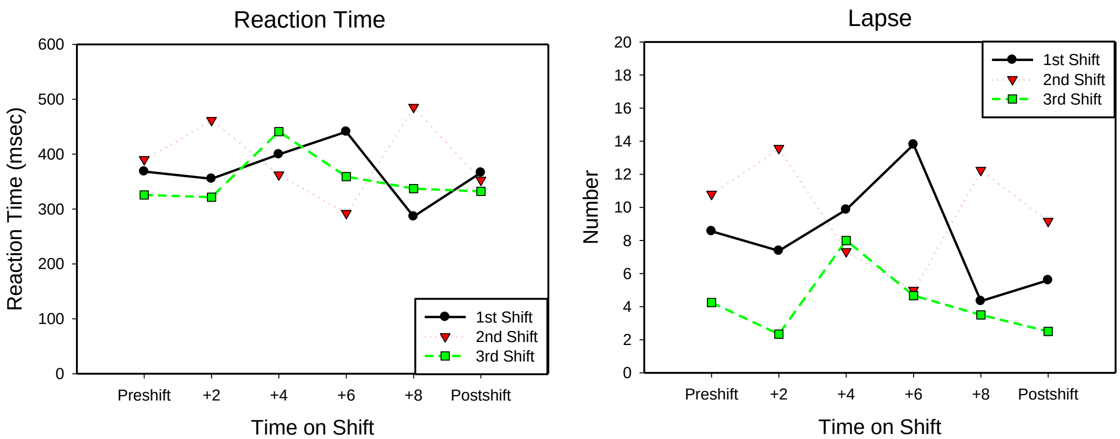

Figure 7. PVT performance by shift and time of day

3.1.5 Wrist activity monitors (WAM) or Actiwatch®**.** The data from the wrist activity monitors (WAM) should be examined separately from the data collected on the PDAs, because some participants who wore the watches did not have data that were included in the other parts of the analysis.

The first shift at WPAFB slept 6 hours on average (366 min), and the second shift obtained almost 6.5 hours on average (383 min). It appears that the third shift got the most sleep with an average of 7 hours and10 minutes (430 min). However, they had the fewest number of sleep periods to contribute to their group average (less than half of what was available for the first shift, and a third of what was available for the second shift), which may have impacted the resulting average. The data are shown in Figure 8 below.

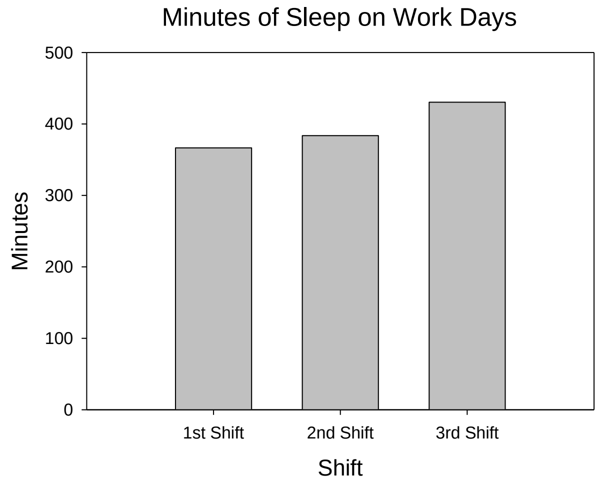

Figure 8. Amount of sleep obtained on work days by shift

3.2 Lackland Air Force Base (LAFB)

As with the WPAFB data set, the number of participants for the LAFB data set was also limited. As with the WPAFB, the means for the reported data are based on the number of tests completed, not the number of participants. Some of the data points are contributed by the same individual over several days of the study. The data points for each shift and time period are presented in Table 3.

Table 3. Number of responses per work shift and time of testing for each variable

Dataset	Shift	Preshift	+2	+4	+6	+8	+10	+12	Postshift
VAS POMS-B PVT	Alpha	17	5	8	7	10	3	5	3
	Bravo	7	4	1	5	3	2	3	2
	Charlie	10	6	7	9	6	3	1	4
	Delta	7	0	6	4	3	5	1	2

Dataset	Shift	n
Sleep Diary	Alpha	16
	Bravo	2
	Charlie	15
	Delta	0
Wrist monitors (WAM)	Alpha	34
	Bravo	27
	Charlie	33
	Delta	31

3.2.1 Visual Analogue Scale (VAS). As with the WPAFB data, direct comparisons among shifts cannot be made statistically due to the low *n*; however, visual comparisons reveal interesting trends. Overall, at LAFB, the levels of alertness, energy, confidence, and sleepiness were all very similar among respondents in the shifts, with levels dropping throughout the work period and sleepiness levels rising.

The individuals in the Bravo shift had the lowest scores for anxious, irritable, and jittery. Those in the Delta shift, reported having slightly higher levels of confidence than those in the rest of the shifts (moderately high vs. moderately low). On the other hand, the Delta shift also reported having slightly higher levels of jittery/nervousness than the other groups.

The participants in the Alpha shift, as a group, reported feeling more energetic and less sleepy than at baseline until about 1515 (10 hours into shift). At 1115 (6 hours into shift), their mean scores for anxiety began to rise, and those levels remained elevated through their tenth hour at work (taken at 1515). At 1315 (8 hours into shift), they also felt more irritable and nervous. By 1515 (10 hours into shift), their levels of irritability returned to baseline, but at this point, the participants in the Alpha shift reported feeling less alert, less energetic, less confident, and more sleepy than at baseline. The results of the end of the day test (12 hours into shift) revealed that the mean scores for all of those feelings (except nervousness) returned to baseline. At post-shift, however, the participants' mean score indicated that the group was once again less alert, energetic and talkative, and more anxious and irritable, but they were also surprisingly less sleepy than at baseline.

Unlike the Alpha shift, the mean scores for the Bravo shift did not indicate that they felt more energetic as the mornings progressed; instead, they reported feeling less alert (2 hours into shift). Their scores for talkativeness and energy also dropped (6 and 8 hours into shift). At 1515 (10 hours into shift), they reported being much less alert and sleepier than they were at baseline; this trend is similar to what was experienced by those in the Alpha shift at the same time of day. However, the participants in the Bravo shift, at the end of the workday (1715), reported feeling more irritable than at baseline, and their levels of sleepiness remained high. Thus, their scores did not return to baseline at the end of the shift like the participants in the Alpha shift had. By the time the Bravo shift had taken the post- shift test, their alertness had dropped even more, and their mean score for energy and confidence were lower. Their reported irritability remained the same as it was at the previous test time, but their group's mean score indicated that they were much sleepier and much more talkative than they were at baseline.

The group sleepiness mean scores for the participants in the Charlie shift stayed close to baseline levels until 2315 (6 hours into shift) when they reported feeling higher levels of sleepiness. Their mean score remained elevated at 0115 (8 hours into shift). At 0315 (10 hours into shift), they reported feeling less energetic, less confident, more nervous, and even sleepier than they were at 2315 and 0115; this change mirrors that which was experienced by the Bravo and Alpha shifts after having worked for 10 hours. Their post-shift mean scores indicated that their sleepiness remained high, and their levels of alertness and energy are lower than they were at baseline.

The group's mean scores for the participants in the Delta shift were much higher at 1915 (2 hours into shift) for anxiety, nervousness, and irritability than at baseline. The group reported feeling less sleepy and less confident as well. By 2115 (4 hours into shift), they indicated feeling less alert, energetic, and talkative. While their level of irritability remained higher than it was at baseline, the score was lower than it had been at 1915. Finally, the group's score for sleepiness became elevated 12 hours into the shift as well. The Delta shift's mean score for anxiety rose from baseline levels at 2315 (6 hours into shift), but it was lower than at 1915 hours. Their sleepiness and nervousness also remained high. At 0115 (8 hours into shift), all of their scores except anxiety returned to baseline levels. Similar to the other three shifts, the participants in the Delta shift reported feeling less alert, talkative, and energetic when they were at 10 hours into their shift (0315). Their reported irritableness remained high, and they indicated that they were much sleepier than they had been at baseline. Scores for nervousness dropped below baseline levels. At the time of the post-shift test, the group means for alertness, energy, and talkativeness were lower than they were at baseline, and their irritability, sleepiness and nervousness were all higher. These results are shown in Figure 9 below.

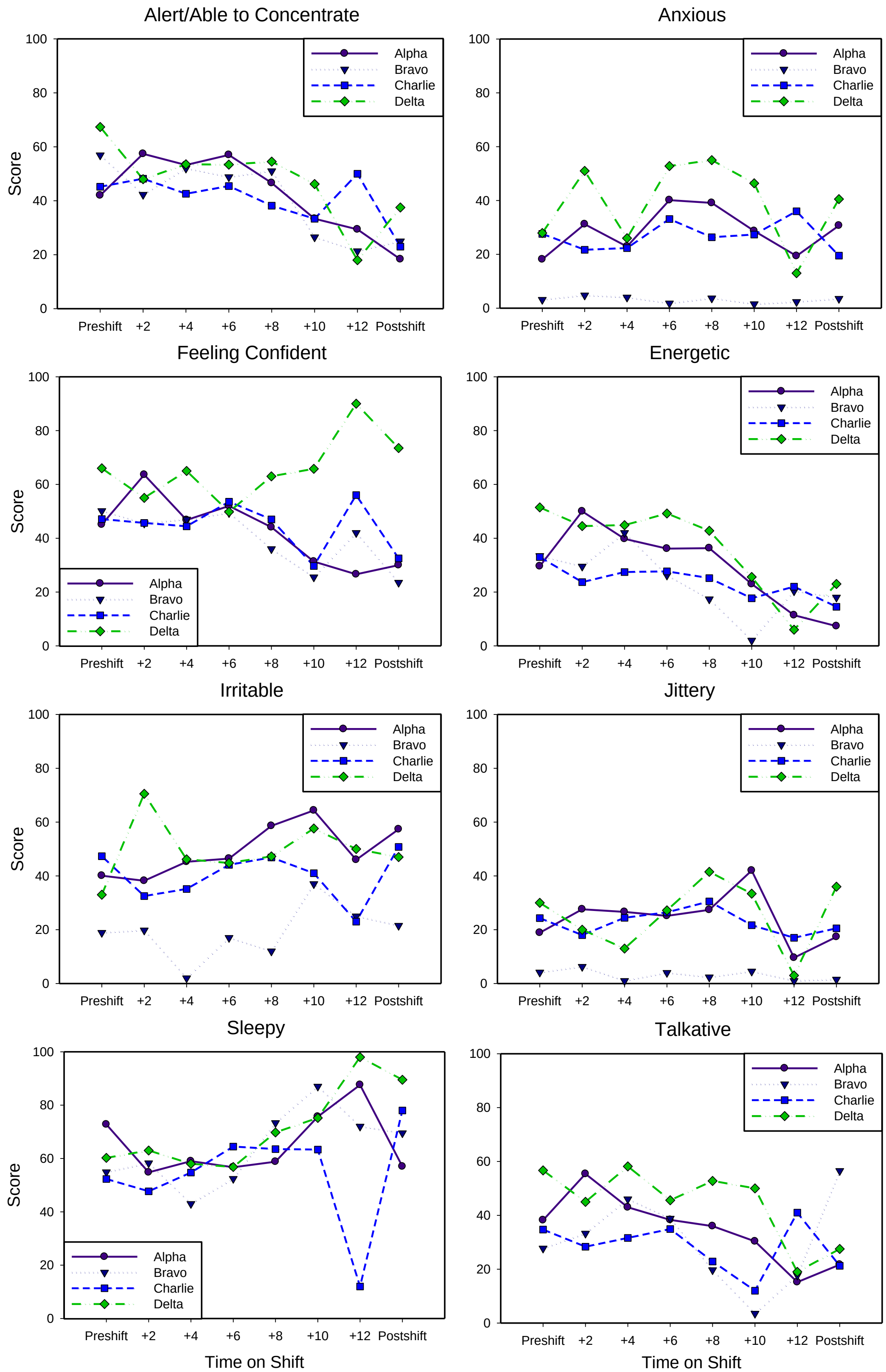

Figure 9. VAS scores by shift and time of day

3.2.2 Profile of Mood States-Brief (POMS-B)

*3.2.2.1 Disturbance Scores***.** The Disturbance Scores at LAFB followed a similar pattern for each of the four shifts. The mean disturbance scores remained close to baseline throughout the work day until the end of the shifts when the reported score rose. For the individuals on Alpha shift, the disturbance rose at 1500, when they had been on duty for 12 hours. For the other three shifts, the participants reported feeling lower moods after they had been on duty for 10 hours. The Bravo and Charlie shifts' mean Disturbance Scores remained elevated until their twelfth hour on duty. All of the groups experienced a return toward their baseline scores at the post-shift test.

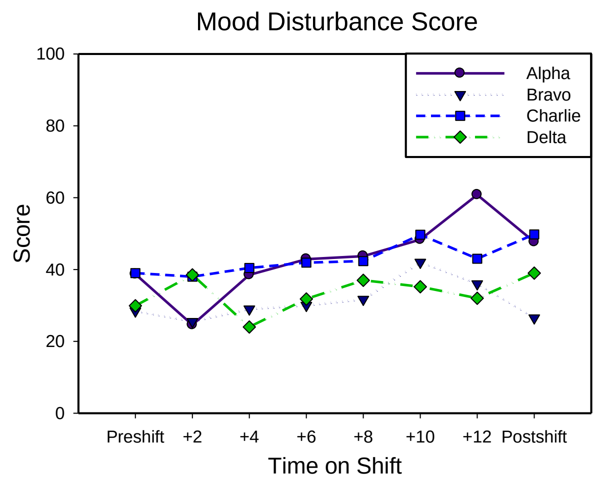

Figure 10. POMS Mood Disturbance Score by shift and time of day

*3.2.2.2 Factor Scores***.** At LAFB, there were some notable changes in the group's mean factor scores throughout their shifts, and many of these changes occurred late in the shifts. Once participants in the Alpha shift had been at work for 12 hours (1715), they reported an increase in tension/anxiety, depression/dejection, anger/hostility, fatigue/inertia, and confusion/ bewilderment compared to baseline. Most of these scores returned to near-baseline levels by the post-shift test; the mean score for tension/anxiety remained elevated. The individuals in the Bravo shift reported increased anger/hostility starting at 1315 (8 hours into shift), and those levels remained high until after their shift ended; the post-shift test revealed mean anger/hostility levels that were below baseline. These participants also reported increased fatigue/inertia beginning at 1515 (10 hours into shift). This mean dropped somewhat at 1715 (12 hours into shift), and then further dropped below baseline levels during the post-shift test. Similarly, the mean score for vigor/activity for the participants in the Charlie shift dropped below baseline at 0315 (10 hours into shift), and their fatigue/inertia score rose. The fatigue/inertia scores remained high at the post-shift testing, but the vigor/activity scores returned to the baseline level. Interestingly, the Charlie shift's mean anger/hostility score rose at post-shift testing. Finally, some of the mean scores for the Delta shift began to change at 0315 (10 hours into shift). Their tension/anxiety, depression dejection, and confusion/bewilderment scores all rose above baseline levels. At 0315 (10 hours into shift), those scores all returned to baseline levels, but the mean score for vigor/activity dropped, and an increase in the fatigue/inertia score. At post-shift, their mean score for vigor/activity decreased even more, and with a higher fatigue/inertia score. Additionally, the mean scores for depression/dejection and confusion bewilderment had both increased from baseline levels.

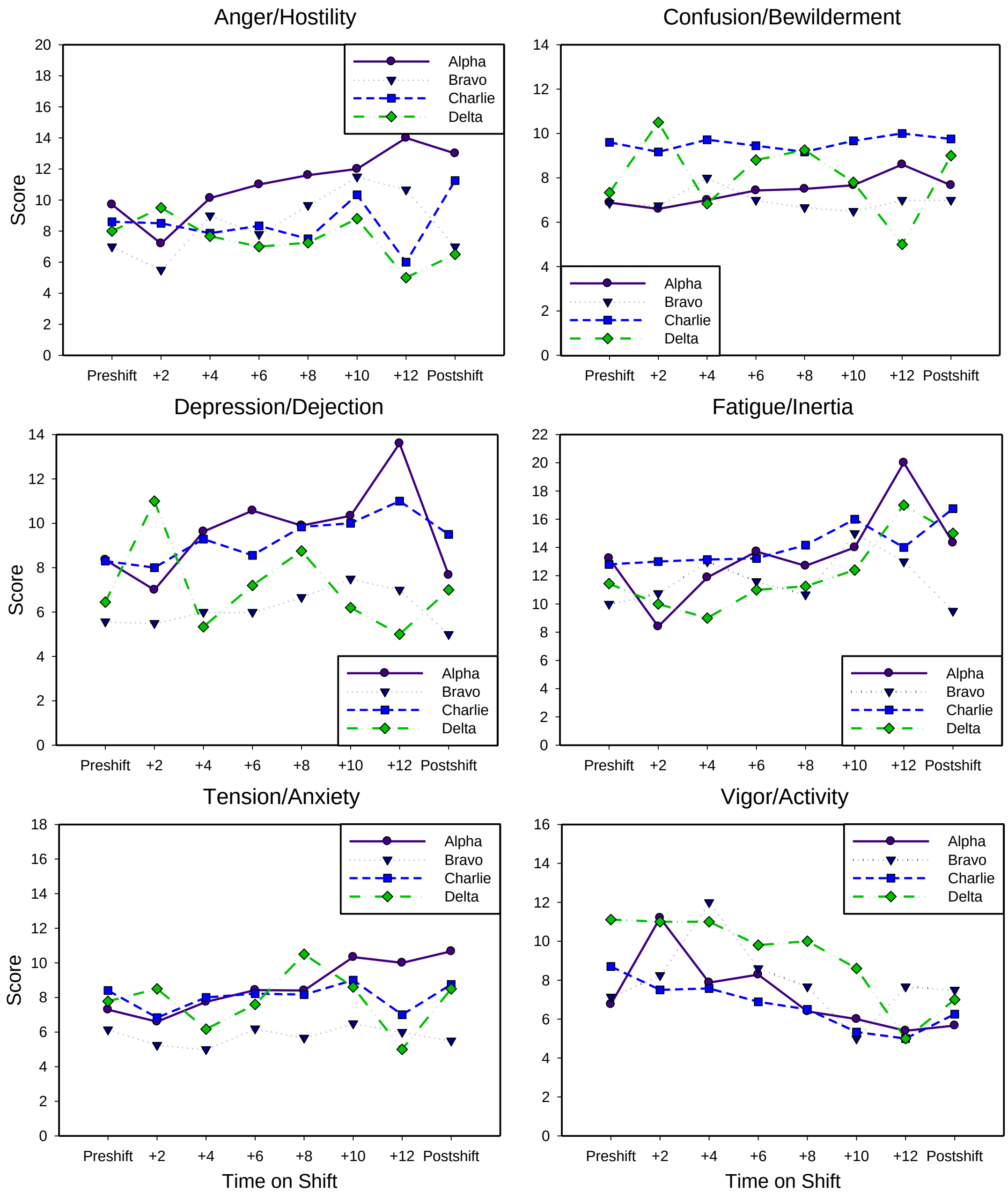

Figure 11. POMS factor scores by shift and time of day

3.2.3 Sleep diary. As mentioned earlier, the *n*'s used for the sleep diaries at both WPAFB and LAFB actually reflect the total number of sleep-periods (that occurred on workdays) from which usable data was pulled, instead of the number of individuals participating. Additionally, only data from the Alpha and Charlie shifts from LAFB will be reported, because no usable data was available from the Delta shift, and the Bravo shift had only two nights of usable data.

The individuals in Charlie shift reported that they slept better, fell asleep easier, slept more deeply, and felt more rested afterwards than did those in the Alpha shift. The group means score for each groups indicated that they both woke around 4 times during the course of their sleep period, and it took both groups roughly 30 minutes to fall asleep at bedtime. (See Figure 12.)

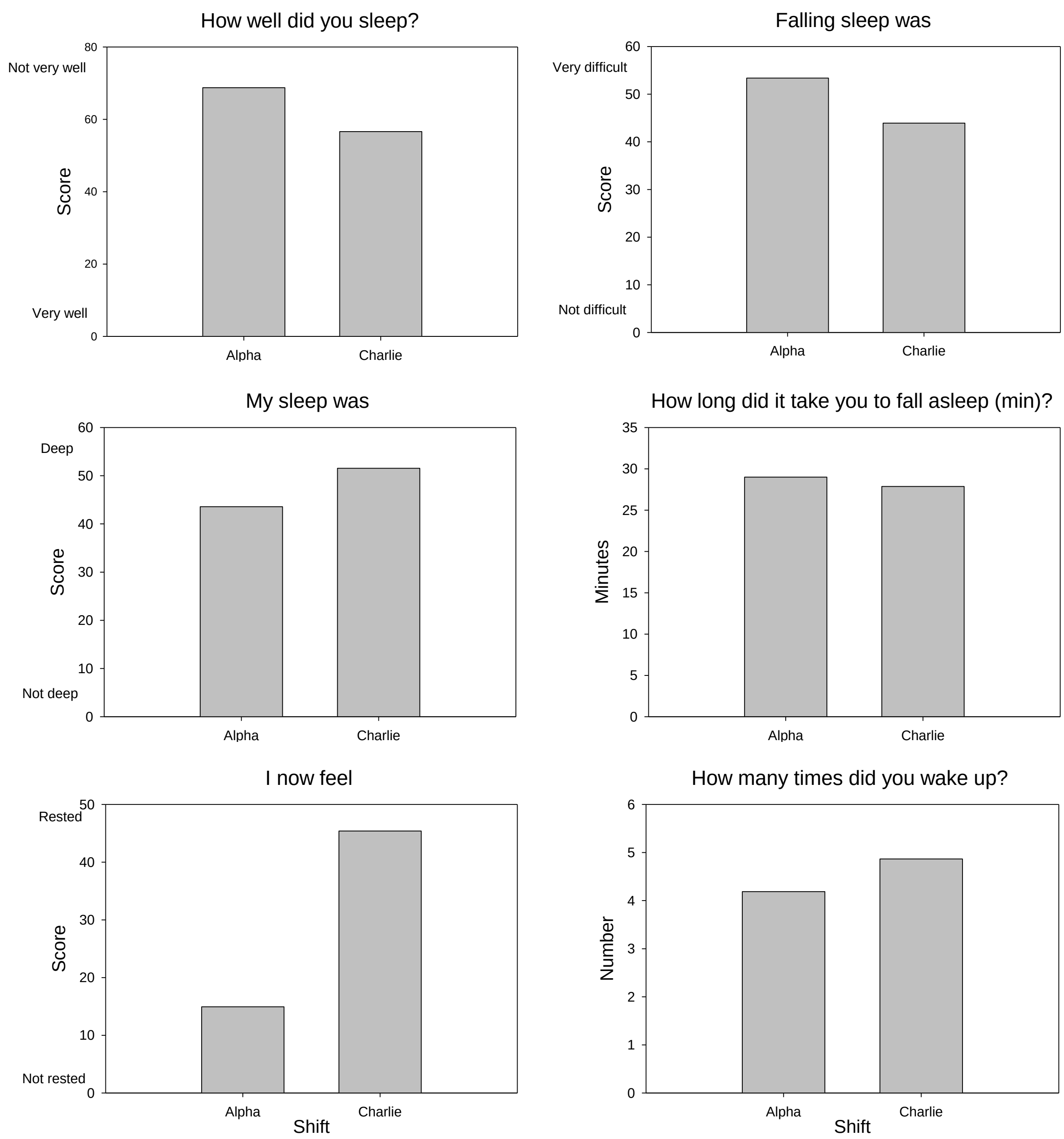

Figure 12. Sleep questionnaire responses concerning sleep quality by shift

Of the times that the groups answered the sleep diary questions, neither group reported any use of over-the-counter sleep aids, herbal remedies, or relaxation techniques to help them sleep. However, both groups reported using good sleep habits while a small percentage of the participants in the Alpha group indicated that they had used alcohol to help them sleep (Figure 13 below).

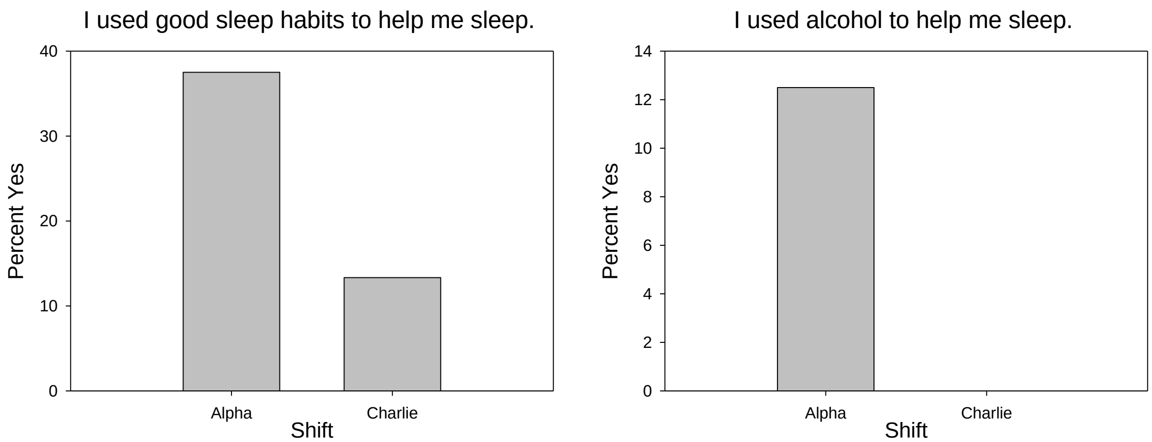

Figure 13. Sleep questionnaire responses by shift concerning aids to obtain sleep

Several questions addressed factors which may have affected the quality of sleep. A percentage of individuals in both groups reported that personal stress/anxiety and not being sleepy affected their sleep; in both cases there were more individuals in the Alpha than in the Charlie group who answered in the affirmative. The data are shown in Figure 14 below.

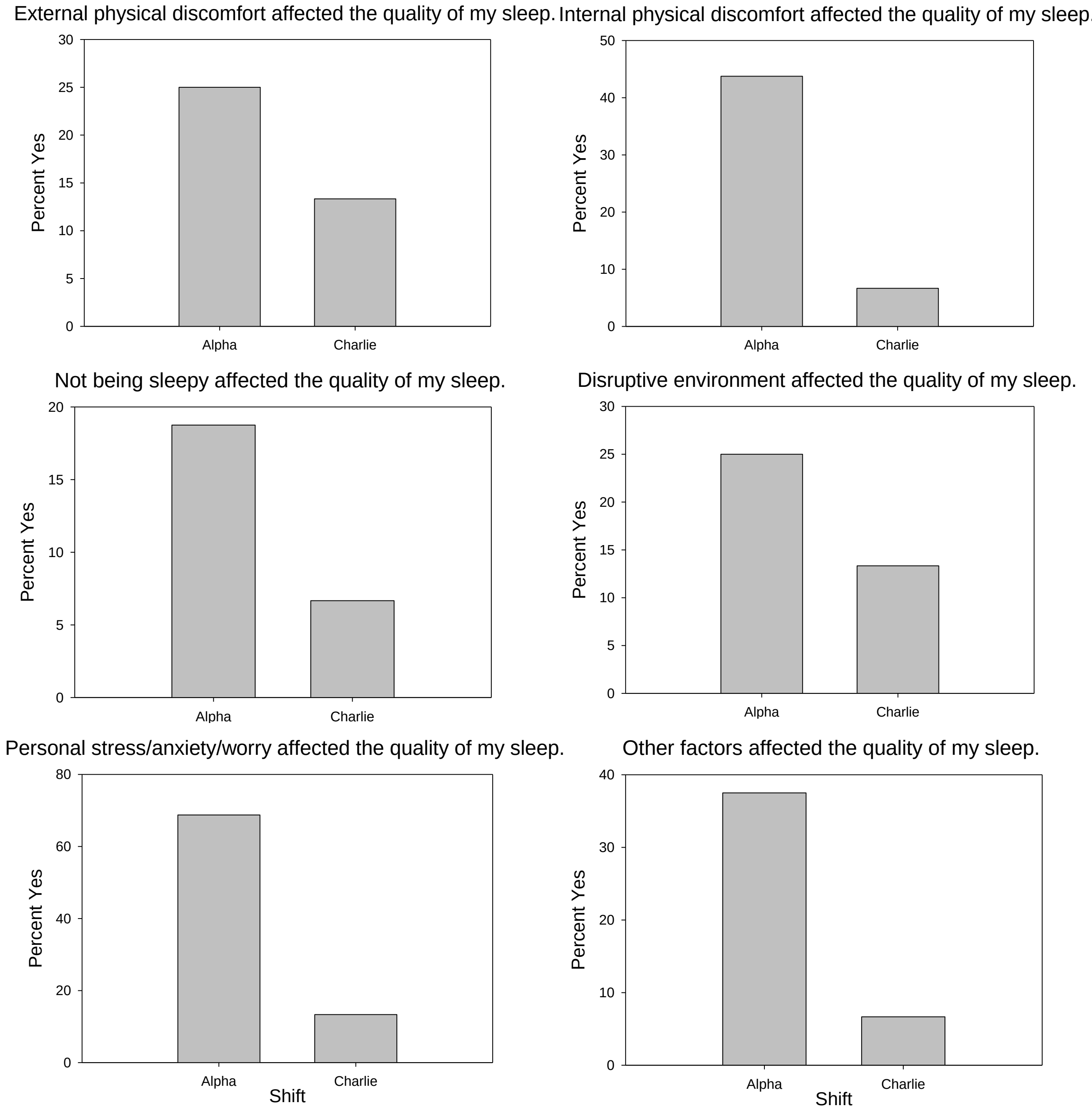

Figure 14. Sleep questionnaire responses by shift concerning factors affecting sleep quality

3.2.4 Psychomotor Vigilance Task (PVT). Performance measures from the PVT indicated longer reaction time and increased number of lapses occurred in the Alpha, Bravo, and Charlie shifts as time on shift increased, particularly after 6 hours on duty. Performance for the participants in the Delta shift showed major changes at the 2-hr on duty test, but nothing after that time. These data are shown in Figure 15 below.

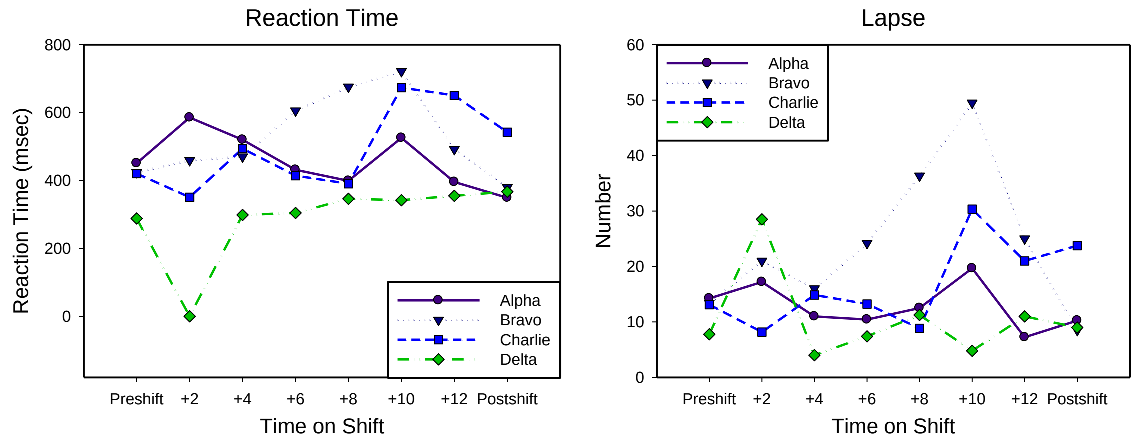

Figure 15. PVT performance by shift and time of day

3.2.5 Wrist activity monitors (WAM) or Actiwatch®**.** As stated previously, the participants who contributed data from the tests (on the PDA) were not necessarily the same participants who wore their wrist activity monitors throughout the study. Therefore, these data should be examined separately from the data collected on the PDAs.

The four shifts at LAFB were all fairly close in the average amounts of sleep they were able to obtain. During the work days, those participants in the Charlie shift attained the most sleep with 406 min. The participants in the Bravo shift had the least amount of sleep on average (351 min). The Alpha and Delta shifts had 381 min and 377 min of sleep, respectively. On their non-work days, the participants from the Delta shift logged the most sleep, closely followed by those in the Alpha and Charlie shifts, with the Bravo shift logging the least amount of sleep. (See Figure 16.)

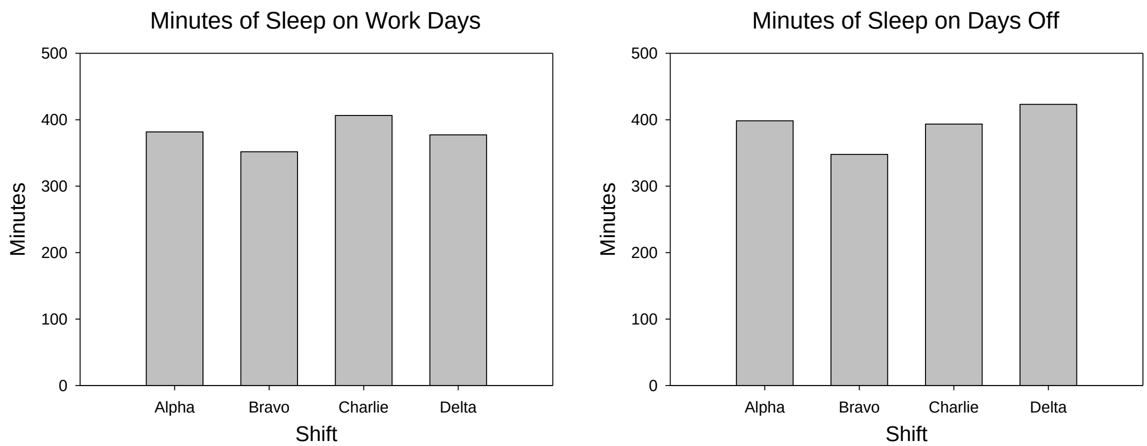

Figure 16. Amount of sleep obtained on work days and days off by shift

4.0 CONCLUSIONS

The present study sought to document the work/rest schedule as well as subjective and objective fatigue levels of Air Force security forces personnel and apply these data to a performance prediction model. Participants in the study included personnel from WPAFB and LAFB, representing two different shift schedules. WPAFB uses 3, 8-hr shifts while LAFB uses 2, 12-hr shifts. Participants were asked to complete a sleep diary each morning and mood and performance tests every 2 hours during their work days, and upon rising, 10 hours into the day, and then prior to bedtime on days off.

Due to the participation variability among shifts and bases, the results from the study do not allow direct comparison of shift schedules nor time on shift as hoped. However, the data were able to show that as time on shift increased, both subjective and objective fatigue generally increased. For example, questions from the POMS and VAS both showed increases in fatigue when the participant was directly asked to describe their levels of sleepiness, fatigue, and alertness. The only objective measure was from the PVT which in some cases showed longer reaction times and increased number of lapses. Caution is used in interpreting these performance data. The PVT is generally performed in a quiet environment while the participant is seated. The present study presented the PVT on a PDA with participants taking it wherever they were at the time the test was scheduled, which may have been a noisy environment and/or while seated or standing.

A subset of the data from LAFB was submitted to MTS Technologies, the contractor for this effort. The data will be used to determine the validation of the SAFTE model as used in the AWARE software.

In addition to validation of the AWARE software, summarized data were presented to the two commanders of the squadrons who participated in the study. While the information will not be used to alter shift schedules, it provided each commander with an idea of how alertness and performance declines across the day, particularly during the night shifts and on the longer, 12-hour shifts.

Due to the participation difficulties in the present study, should the security forces command desire to investigate their shift schedules for fatigue effects, participants in a future study should be screened carefully for commitment to the study so consistent data can be obtained to make better conclusions regarding mood and performance. In addition, a true comparison of 8- versus 12-hr shifts will require consistent data collection methodologies.

REFERENCES

Åkerstedt T. (1995). Work hours, sleepiness and the underlying mechanisms. *Journal of Sleep Research*, 4, 15-22.

Balkin, T.J., Rupp, T., Picchione, D., and Wesensten, N.J. (2008). Sleep loss and sleepiness: Current issues. *Chest*, 134, 653-660.

Belenky, G., Wesensten, N.J., Thorne, D.R., Thomas, M.L., Sing, H.C., Redmond, D.P., Russo, M.B., and Balkin, T.J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study. *Journal of Sleep Research,* 12, 1-12.

Caldwell, J.A., Caldwell, J.L, and Schmidt, R.M. (2008) Alertness management strategies for operational contexts. *Sleep Medicine Reviews*, 12, 257-273.

Coleman, R.M., and Dement, W.C. (1986). Falling asleep at work: A problem for continuous operations. *Sleep Research*, 15, 265.

Dinges, D. (1994). An overview of sleepiness and accidents. *Journal of Sleep Research,* 4(suppl 2), 4-14.

Hursh, S.R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., Storm, W. F., Miller, J. C., and Eddy, D. R. (2004). Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine*, 75(3 Suppl), A44-53; discussion A54-60.

Krueger, G.P. (1989). Sustaining military performance in continuous operations: Combatant fatigue, rest and sleep needs. *Handbook of military psychology*. New York: John Wiley and Sons, pp 255-277.

National Transportation Safety Board. (1990) Marine accident report-grounding of the U.S. tankship EXXON VALDEZ on Bligh Reef, Prince William Sound, near Valdez, Alaska, March 24, 1989. Washington, DC: National Transportation Safety Board; 1990. NTSB/MAR-90/04.

Philip, P., and Akerstedt, T. (2006). Transport and industrial safety, how are they affected by sleepiness and sleep restriction? *Sleep Medicine Review*, 10(5), 347-356.

Torsvall, L. and Akerstedt, T. (1987). Sleepiness on the job: Continuously measured EEG changes in train drivers. *Electroencephalography and Clininical Neurophysiology*, 66, 502-11.

Torsvall, L., Akerstedt, T., Gillander, K., and Knutsson, A. (1989). Sleep on the night shift: 24-hour EEG monitoring of spontaneous sleep/wake behavior. *Psychophysiology*, 26, 352-8.

U.S. Nuclear Regulatory Commission. (1987). Report on the accident at the Chernobyl Nuclear Power Station. Washington, DC: US Government Printing Office.

Van Dongen, H. P., Maislin, G., Mullington, J. M., and Dinges, D. F. (2003). The cumulative cost of additional wakefulness: does-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep, 26*(2), 117-26.

ABBREVIATIONS

AFB	Air Force Base
DoD	Department of Defense
LAFB	Lackland Air Force Base
PDA	Personal digital assistant
WPAFB	Wright Patterson Air Force Base

ACRONYMS

AWARE	Aggregate Wakefulness And Readiness Estimator
msec	millisecond
POMS-B	Profile of Mood States – Brief
PVT	Psychomotor Vigilance Test
SAFTE	Sleep, Activity, Fatigue, and Task Effectiveness
VAS	Visual Analogue Scale
WAM	Wrist activity monitor